What is claimed is:

1. A method for pictographic recognition, comprising:

isolating words in a scanned document image;

generating image graphs for each isolated word;

obtaining relevant search image graphs and corresponding search isomorphic keys generated from sample documents;

comparing the geometry of the isolated word image graphs to the geometry of the search image graphs; and

matching the isolated word image graphs with the search image graphs using the search isomorphic keys.

2. The method of claim 1, further comprising providing search terms and obtaining relevant search image graphs and corresponding search isomorphic keys generated from sample documents based on the search terms provided.

3. The method of claim 1, further comprising segmenting the isolated word image graphs, and wherein matching the isolated word image graphs with the search image graphs comprises matching segments within the isolated word image graphs to segments within the search image graphs.

4. The method of claim 3, wherein segmenting the isolated word image graphs comprises finding a baseline path through each isolated word image graph and segmenting the isolated word image graph based on certain conditions.

5. The method of claim 3, wherein matching segments comprises using a flexible window to look at segments within each isolated word image graph and to determine if the geometry of a segment in an isolated word image graph matches the geometry of a segment in a search image graph.

6. The method of claim 5, wherein determining if a segment in an isolated word image graph matches a segment in a search image graph comprises generating a isomorphic key for each segment in an isolated word image graph and comparing the isolated word isomorphic keys to the search isomorphic keys.

7. The method of claim 6, wherein determining if a segment in an isolated word image graph matches a segment in a search image graph further comprises directly matching the geometry of a matched sample word image graph to the geometry of the matched search image graph using feature measurements.

8. The method of claim 1, further comprising storing matched isolated word image graphs in a result matrix, and assigning a confidence factor to the image graphs stored in the results matrix.

9. The method of claim 8, wherein assigning a confidence factor comprises extracting Tuples from the matched isolated word image graphs stored in the results matrix and matching the extracted Tuples against Tuples extracted form the search image graphs.

10. The method of claim 9, further comprising assigning a score to the isolated word image graphs based on the degree of matching determined by

comparing Tuples for the isolated word image graphs to Tuples generated for the search image graphs, and flagging sample word image graphs with a score above a certain threshold for further review.

11. The method of claim 1, further comprising generating a source language library, wherein generating a source language library comprises scanning sample documents that are in the source language, isolating words in the scanned documents, generating search image graphs for the words isolated in the sample documents, and generating search isomorphic keys corresponding to the words isolated in the sample documents.

12. The method of claim 11, wherein the search image graphs and search isomorphic keys are obtained from the source language library.

13. The method of claim 11, wherein the search terms are generated for a target language.

14. A method for pictographic recognition, comprising:

generating image graphs for words in a document;

obtaining search image graphs;

comparing the structural patterns of the generated image graphs with the structural patterns of the search image graphs; and

identifying certain words in the document based on the comparison of the structural patterns.

15. The method of claim 14, wherein generating image graphs for words in a document comprises:

scanning the document to generate a document image;

isolating words in the scanned document image; and

generating image graphs for each isolated word.

16. The method of claim 15, wherein comparing structural patterns comprises:

obtaining relevant search image graphs and corresponding search isomorphic keys generated from sample documents;

comparing the geometry of the isolated word image graphs to the geometry of the search image graphs; and

matching the isolated word image graphs with the search image graphs using the search isomorphic keys.

17. The method of claim 16, further comprising providing search terms and obtaining the search image graphs and corresponding search isomorphic keys based on the search terms provided.

18. The method of claim 16, further comprising segmenting the isolated word image graphs; and wherein matching the isolated word image graphs with the search image graphs comprises matching segments within the isolated word image graphs to segments within the search image graphs.

19. The method of claim 18, wherein segmenting the isolated word image graphs comprises finding a baseline path through each isolated word image graph and segmenting the isolated word image graph based on certain conditions.

20. The method of claim 18, wherein matching segments comprises using a flexible window to look at segments within each isolated word image graph and to determine if the geometry of a segment in an isolated word image graph matches the geometry of a segment in a search image graph.

21. The method of claim 20, wherein determining if a segment in an isolated word image graph matches a segment in a search image graph comprises generating a isomorphic key for each segment in an isolated word image graph and comparing the isolated word isomorphic keys to the search isomorphic keys.

22. The method of claim 21, wherein determining if a segment in an isolated word image graph matches a segment in a search image graph further comprises directly matching the geometry of a matched sample word image graph to the geometry of the matched search image graph using feature measurements.

23. The method of claim 16, further comprising storing matched isolated word image graphs in a result matrix, and assigning a confidence factor to the image graphs stored in the results matrix.

24. The method of claim 23, wherein assigning a confidence factor comprises extracting Tuples from the matched isolated word image graphs stored in the results matrix and matching the extracted Tuples against Tuples extracted from the search image graphs.

25. The method of claim 24, further comprising assigning a score to the isolated word image graphs based on the degree of matching determined by comparing Tuples for the isolated word image graphs to Tuples generated for the

search image graphs, and flagging sample word image graphs with a score above a certain threshold for further review.

26. An image recognition system for searching documents in a source language comprising:

an imaged document, the imaged document being stored in a document database;

a source language library for storing source language search terms;

an image graph constructor coupled to the document database and the source language library, the image graph constructor configured to,

generate search term image graphs from the source language search terms, and

generate a collection of image graphs representing the imaged document;

an image graph database for storing the search term image graphs and the collection of image graphs generated by the image graph constructor; and

a comparison module coupled to the image graph database, the comparison module configured to search the imaged documents by comparing the collection of image graphs with selected search term image graphs;

wherein if at least one image graph from the collection of image graphs matches the selected search term image graphs, the imaged document is flagged as containing a search term justifying further analysis of the document.

27. The system of claim 26, wherein the source language is any language foreign to a user of the system.

28. The system of claim 26, wherein an imaged document is an electronic version of a physical source language document.

29. The system of claim 28, wherein the imaged document is generated by scanning the physical source language document.

30. The system of claim 26, wherein the source library further includes a collection of source language characters.

31. The system of claim 30, wherein the collection of source language characters includes written and typographical variations of each source language character.

32. The system of claim 26, wherein the image graph constructor is generates an image graph for each character contained in the source language library.

33. The system of claim 26, wherein the image graph constructor is configured to employ an image graph generation process.

34. The system of claim 33, wherein the image graph generation process involves an image reduction process and a data storage process.

35. The system of claim 34, wherein the image reduction process involves extracting information concerning the character's link and node connections.

36. The system of claim 33, wherein the image graph constructor is configured to perform an image graph generation process, the process comprising the steps of:

reducing the character to a skeleton image;

representing the skeleton image of the character in the form of a linked list comprising a plurality of entries and a plurality of pointers between the entries, organized on the basis of internal structure corresponding to a plurality of nodes, and connections between the plurality of nodes, wherein each of the plurality of entries in the linked list corresponds to one of the plurality of nodes, and each of the pointers between entries corresponds to one of the connections between nodes; and

storing the image graph of the character as the representation of the internal structure of the character.

37. The system of claim 34, wherein the data storage process stored the skeleton image of character as a data structure having a header and a connectivity network.

38. The system of claim 34, wherein the header includes at least a Connectivity Array and a Connectivity Key.

39. The system of claim 38, wherein the reference for a given node of the character being a series of integers, each integer being equal to the number of nodes having a given number of connections between nodes, connected to the given node

40. The system of claim 38, wherein the Connectivity Key is unique for a given plurality of nodes and connections between the given plurality of nodes, the unique Connectivity Key corresponding to a Connectivity Array for each node of the character.

41. The system of claim 26, wherein the imaged document stored in the document database and input into the image graph constructor is transformed into a collection of image graph by the image graph generation process.

42. The system of claim 41, wherein the image reduction process and the data storage process is executed on the imaged document to produce a collection of image graphs.

43. The system of claim 26, wherein the comparison module is configured to execute a screening process and a searching process.

44. The system of claim 43, wherein the screening process conducts two screening methods including a screen by Connectivity Key method and a screen by Connectivity Array method.

45. The system of claim 44, wherein the screening by Connectivity Key is conducted to determine the Connectivity Key contained in the header of the imaged graph search term match the Connectivity Key contained in the header of an image graph of the collection of image graphs.

46. The system of claim 45, wherein if the screening by Connectivity Key is successful, the searching process is activated.

47. The system of claim 44, wherein if the screening by Connectivity Array is conducted by comparing the Connectivity Array associated with the search term image graphs with the each Connectivity Array associated with the collection of image graphs.

48. The system of claim 45, wherein if the screening by Connectivity Array is successful, the searching process is activated.

49. The system of claim 43, wherein the searching process employs a depth first search.

50. The system of claim 43, wherein the searching process computes the number of connections between the nodes for each image graph in the collection of image graphs to the number of connection nodes in the search term image graph for at least first and second ones of the search term image graphs and link ratios for the first and second stored search term image graphs.

51. A method for identifying search terms in a foreign language document comprising:

establishing a foreign language lexicon, the foreign language lexicon including foreign language search terms and foreign language characters;

providing an imaged document in a foreign language;

inputting the foreign language search terms and the imaged document into an image graph constructor, the image graph constructor configured to generate search term image graphs and configured to generate a collection of image graphs representing the imaged documents;

searching the imaged documents by comparing the collection of image graphs to selected search term image graphs; and

flagging imaged documents when an image graph from the collection of image graphs matches the selected search term image graphs.

52. The method of claim 51, the establishment of the foreign language lexicon comprising:

identifying characters associated with the foreign language; and

identifying written and typographical variations for each character.

53. The method of claim 51, the establishment of the foreign language lexicon further comprising:

identifying search terms in a domestic language; and

interpreting the search term from the domestic language to the foreign language using an interpreter fluent in the foreign language.

54. The method of claim 51, wherein the foreign language lexicon includes foreign language characters.

55. The method of claim 53, wherein the interpreter supplements the foreign language lexicon by identifying, in the foreign language all colloquial variations of the search terms in the foreign language.

56. The method of claim 51, wherein the imaged document is an electronic version of a physical foreign language document.

57. The method of claim 56, wherein the imaged document is generated by scanning the physical foreign language document.

58. The method of claim 51, wherein the image graph constructor generates an image graph for each character contained in the foreign language lexicon

59. The method of claim 51, wherein the image graph constructor is configured to employ an image graph generation process.

60. The method of claim 59, wherein the image graph generation process involves an image reduction process and a data storage process.

61. The method of claim 60, wherein the image reduction process involves extracting information concerning the character's link and node connections.

62. The method of claim 61, the image graph constructor is configured to perform an image graph generation process, the process comprising the steps of:

reducing the character to a skeleton image;

representing the skeleton image of the character in the form of a linked list comprising a plurality of entries and a plurality of pointers between the entries, organized on the basis of internal structure corresponding to a plurality of nodes, and connections between the plurality of nodes, wherein each of the plurality of entries in the linked list corresponds to one of the plurality of

nodes, and each of the pointers between entries corresponds to one of the connections between nodes; and

storing the image graph of the character as the representation of the internal structure of the character.

63. The method of claim 60, wherein the data storage process stores the skeleton image of character as a data structure having a header and a connectivity network.

64. The method of claim 63, wherein the header includes at least a Connectivity Array and a Connectivity Key.

65. The method of claim 64, wherein the reference for a given node of the character being a series of integers, each integer being equal to the number of nodes having a given number of connections between nodes, connected to the given node

66. The method of claim 64, wherein the Connectivity Key is unique for a given plurality of nodes and connections between the given plurality of nodes, the unique Connectivity Key corresponding to a Connectivity Array for each node of the character.

67. The method of claim 51, wherein the imaged document stored in the document database and input into the image graph constructor is transformed into a collection of image graph by the image graph generation process.

68. The method of claim 67, wherein the image reduction process and the data storage process is executed on the imaged document to produce a collection of image graphs.

69. The method of claim 51, wherein the comparison module is configured to execute a screening process and a searching process.

70. The method of claim 68, wherein the screening process conducts two screening methods including a screen by Connectivity Key method and a screen by Connectivity Array method.

71. The method of claim 70, wherein the screening by Connectivity Key is conducted to determine the Connectivity Key contained in the header of the imaged graph search term match the Connectivity Key contained in the header of an image graph of the collection of image graphs.

72. The method of claim 71, wherein if the screening by Connectivity Key is successful, the searching process is activated.

73. The method of claim 71, wherein if the screening by Connectivity Array is conducted by comparing the Connectivity Array associated with the search term image graphs with the each Connectivity Array associated with the collection of image graphs.

74. The method of claim 71, wherein if the screening by Connectivity Array is successful, the searching process is activated.

75. The method of claim 68, wherein the searching process employs a depth first search.

76. The method of claim 68, wherein the searching process computes the number of connections between the nodes for each image graph in the collection of image graphs to the number of connection nodes in the search term image graph for at least first and second ones of the search term image graphs and link ratios for the first and second stored search term image graphs.

77. The method of claim 51, the image graph generation step further including:

creating an image of each of the characters contained in the foreign language lexicon;

reducing the image of the character to a skeleton image;

ordering the plurality of nodes in a first order;

reordering the plurality of nodes in a second order, wherein the second order is the same of all characters having the same number of nodes connected by the same number of connections between nodes in the same manner as the test character;

representing the skeleton image of the character on the basis of the internal structure of the character by a descriptor corresponding to a plurality of nodes and connections between the plurality of nodes of the character, wherein the descriptor is a Connectivity Key which unique for a given plurality of nodes and connections between the given plurality of nodes, the unique

Connectivity Key corresponding to the second order of the plurality of nodes; and

storing the Connectivity Key representation of the internal structure of the character as the descriptor of the character.